# 1.3 凸集、凸函数、凸规划

## 1.3.1 凸集

定义1.3.1 给定非空集合$F \subseteq R^n$,如果$\forall x, y \in F, \alpha \in [0,1]$都有

$$\alpha x + (1-\alpha) y \in F$$ 称为x与y的凸组合

那么称$F$为$R^n$中的一个凸集。

如果凸集为开集，则称为开凸集；若凸集为闭集，则称为闭凸集。

规定：空集为凸集；

注：单点集与$R^n$为凸集

例1.3.1 假设$\omega \in R^n \setminus \{0\}, \beta \in R$,证明下面的集合为凸集.

1).超平面$H = \{x \in R^n \mid \omega^T x = \beta\}$;

2).闭半空间$\{x \in R^n | \omega^T x \geq \beta\}$和 $\{x \in R^n | \omega^T x \leq \beta\}$;

3). 开半空间$\{x \in R^n \mid \omega^T x > \beta\}$ 和$\{x \in R^n \mid \omega^T x < \beta\}$;

4). 超球 $B = \{x \in R^n \mid \|x\| \leq \beta\}$,其中$\beta \geq 0$.

证1) $\forall x, y \in H$,有$\omega^T x = \beta$, $\omega^T x = \beta$.对$\forall \alpha \in [0,1]$,有

$$\omega^T(\alpha x + (1-\alpha)y) = \alpha\omega^T x + (1-\alpha)\omega^T y$$

$$= \alpha\beta + (1-\alpha)\beta = \beta,$$

从而$\alpha x + (1-\alpha)y \in H$, 故超平面$H$为凸集.

## 命题1.3.1

假设$F_i \subseteq R^n$为凸集且$\beta_i \in R$(其中$i \in \{1,2,\cdots,p\}$)，则下列集合均为凸集

1)交集　$F := F_1 \cap F_2 \cap \cdots \cap F_p$

2) 集合　$\beta_i F_i := \{\beta_i x \mid x \in F_i\}$;

3)和集　$F_1 + F_2 := \{x + y \mid x \in F_1, y \in F_2\}$;

4)集合　$\sum_{i=1}^{p} \beta_i F_i$ .

证明： 1) 任取$x, y \in F$,则$x, y \in F_i \ (i=1,2,\cdots,p)$.

对$\forall \alpha \in [0,1]$，因为$F_i$是凸集, 因此有 $\alpha x+(1-\alpha)y \in F_i \ (i=1,2,\cdots,p)$

故 $\alpha x+(1-\alpha)y \in F_1 \cap F_2 \cap \cdots \cap F_p = F$,

所以$F$为凸集.

定理1.3.1 非空集合$F \subseteq R^n$为凸集$\Leftrightarrow$对$\forall x_i \in F$及任意满足$\sum_{\mathrm{i}}^{p} \alpha_{\mathrm{i}} = 1$的非负实数$\alpha_i (i \in \{1, 2, \cdots p\}$，且$p \geq 2$的正整数)，都有$\sum_{i=1}^{p} \alpha_i x_i \in F$.

证明：由凸集定义得定理的充分性。下用归纳法证明必要性。

$p = 2$时由定义知结论成立；

假设 $p = k$时结论成立，下证 $p = k + 1$时结论成立。

对$\forall x^i \in F$，以及满足$\sum_{i=1}^{k+1} \alpha_i = 1$的非负实数$\alpha_i (i \in \{1, 2, \cdots, k, k+1\})$,有

$$\sum_{i=1}^{k+1}\alpha_i x^i=(\sum_{i=1}^{k}\alpha_i x^i)+\alpha_{k+1}x^{k+1}$$

$$=(1-\alpha_{k+1})(\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_{k+1}}x^i)+\alpha_{k+1}x^{k+1}.$$

由$\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_{k+1}}=1,\frac{\alpha_i}{1-\alpha_{k+1}}\geq 0$及归纳假设知$\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_{k+1}}x^i\in F$,

再根据凸集的定义有 $(1-\alpha_{k+1})(\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_{k+1}}x^i)+\alpha_{k+1}x^{k+1}\in F$,

即 $\sum_{i=1}^{k+1}\alpha_i x^i\in F$. 由归纳法原理知结论成立.

## 定义1.3.2

假设$F_1, F_2 \subseteq R^n$为两个非空凸集. 如果存在非零向量$\omega \in R^n$和实数$t$, 使得

1)对$\forall x \in F_1, y \in F_2$都有$\omega^{\mathrm{T}} x \geq t$且$\omega^{\mathrm{T}} y \leq t$,则称超平面

$$\pi := \{x \in R^n | \omega^T x = t\}$$ 分离凸集$F_1$和$F_2$;

2)对$\forall x \in F_1, y \in F_2$都有$\omega^{\mathrm{T}} x > t$且$\omega^{\mathrm{T}} y < t$,则称超平面

$$\pi := \{x \in R^n | \omega^T x = t\}$$ 严格分离凸集$F_1$和$F_2$;

## 定理1.3.2 （点与凸集分离定理）

设$F$为$R^n$中的非空闭凸集，$x^0 \in R^n$且$x^0 \notin F$,则存在$R^n$中的超平面严格分离集合$F$和点$\{x^0\}$.

证明：记$\Omega:=\{y \mid y = x - x^0, x \in F\}$. 显然 $0 \notin \Omega$.

对任意$y^1, y^2 \in \Omega$,存在$x^1, x^2 \in F$,使得 $y^1 = x^1 - x^0, y^2 = x^2 - x^0$.

从而对$\forall \alpha \in [0,1], \alpha y^1 + (1-\alpha) y^2 = \alpha x^1 + (1-\alpha) x^2 - x^0$

由$F$为凸集知$\alpha x^1 + (1-\alpha) x^2 \in F$,因而$\alpha y^1 + (1-\alpha) y^2 \in \Omega$，故$\Omega$为凸集.

对$\forall \delta > 0$, 记$N_\delta(0) := \{\mathrm{y} \in R^n \mid \|y\|_2 \le \delta\}$.

选取适当的$\delta > 0$,使得$N_\delta(0) \cap \Omega$为非空有界闭集，

所以连续函数$\|y\|_2$在$N_\delta(0) \cap \Omega$可达到最小值, 记其为$\hat{y}\ (\in \Omega)$,

则对任意 $y \in \Omega$ 都有$\|\hat{y}\|_2 \le \|y\|_2$.

因此对$\forall \varepsilon \in [0,1], \forall y \in \Omega, \|\hat{y}\|_2 \le \|\varepsilon y + (1-\varepsilon)\hat{y}\|_2$,即

$$\hat{y}^T \hat{y} \le (\varepsilon y + (1-\varepsilon)\hat{y})^T (\varepsilon y + (1-\varepsilon)\hat{y}) = \varepsilon^2 y^T y + 2\varepsilon(1-\varepsilon)\hat{y}^T y + (1-\varepsilon)^2 \hat{y}^T \hat{y}$$

$$= \varepsilon^2 (y^T y - 2\hat{y}^T y + \hat{y}^T \hat{y}) + 2\varepsilon \hat{y}^T y + \hat{y}^T \hat{y} - 2\varepsilon \hat{y}^T \hat{y}$$

$$= \varepsilon^2 \|y - \hat{y}\|^2 + 2\varepsilon (y - \hat{y})^T \hat{y} + \hat{y}^T \hat{y}.$$

从而$\varepsilon^2\|y-\hat{y}\|_2^2+2\varepsilon(y-\hat{y})^T\hat{y}\ge 0$.

由$\varepsilon\in[0,1]$的任意性可得$(y-\hat{y})^T\hat{y}\ge 0$，即$y^T\hat{y}\ge\|\hat{y}\|_2^2$.

对$\forall x\in F$,记 $y:=x-x^0(\in\Omega)$,则对$\forall x\in F$,都有$(x-x^0)^T\hat{y}\ge\|\hat{y}\|_2^2$,

即$x^T\hat{y}\ge(x^0)^T\hat{y}+\|\hat{y}\|_2^2$.

令$\omega=\hat{y},\ t=(x^0)^T\hat{y}+\dfrac{1}{2}\|\hat{y}\|_2^2$,则$\omega^{\mathrm{T}}x>t,\omega^T x^0<t$.

因此超平面$\pi:=\{\mathrm{x}\in R^n:\omega^T x=t\}$可严格分离凸集$F$与点$x^0$.

## 引理1.3.1（Farkas引理）

设$A \in R^{m\times n}, b \in R^n$,则不等式组 $Ax \le 0,\ b^T x > 0$， (1)

与不等式组 $A^T y = b,\ y \ge 0$ (2)

有且仅有一组有解.

证明：假设(2)有解，即存在$y \in R^m$使得$A^T y = b, y \ge 0$.

若(1)有解, 即存在$x \in R^n$使得$Ax \le 0$,则

$b^T x = (A^T y)^T x = y^T Ax \le 0.$

从而若(2)有解，则(1)必无解.

若(2)无解，下证(1)有解.

若(2)无解，下证(1)有解.

记$\Omega := \left\{ z := A^T y \mid y \geq 0 \right\}$，则$\Omega \subseteq R^n$为非空闭凸集且$b \notin \Omega$.

由定理1.3.2知，存在$\omega \in R^n, t \in R$使得

$\omega^T b > t, \omega^T z < t, \forall z \in \Omega$. 由$0 \in \Omega$知$t > 0$.

$t > \omega^T z = \omega^T (A^T y) = y^T A\omega, y \geq 0$.

由$y$中分量可以任意大可知，$A\omega \leq 0$; 又$b^T w > t > 0$, 从而$\omega$为(1)的解.

## 定理1.3.3

设$p$和$q$是两个非负整数，$u^0, u^1, \cdots, u^p, v^1, \cdots . v^q \in R^n$,则等式与不等式组

$$d^T u^0 < 0,\ \ d^T u^i = 0(i \in \{1, 2, \cdots, p\}),\ \ d^T v^i \geq 0(i \in \{1, 2, \cdots, q\})(1)$$

无解

$\Leftrightarrow$ 存在实数$\alpha_i (i \in \{1, 2, \cdots, p\})$和非负实数$\beta_i (i \in \{1, 2, \cdots, \})$使得;

$$u^0 = \sum_{i=1}^{p} \alpha_i u^i + \sum_{i=1}^{q} \beta_i v^i (2)$$

证明：$d^T u^i = 0 \Leftrightarrow d^T u^i \geq 0, (-d)^T u^i \geq 0, (i = 1, 2, \cdots, p)$;

令$A = (-u^1, \cdots, -u^p, u^1, \cdots, u^p, -v^1, \cdots, -v^q)^T$,则（1）可写成

$Ad \leq 0, (-u^0)^T d > 0.$

由$Farkas$引理,(1)无解 $\Leftrightarrow$ 存在$(x, y, z) \in R_+^p \times R_+^p \times R_+^q$使得

$$-u^0 = A^T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sum_{i=1}^{p} x_i(-u^i) + \sum_{i=1}^{p} y_i u^i + \sum_{i=1}^{q} z_i(-v^i).$$

令$\alpha_i := x_i - y_i (i = 1, 2, \cdots, p), \beta_i := z_i \geq 0 (i = 1, \cdots, q)$则

$$u^0 = \sum_{i=1}^{p} \alpha_i u^i + \sum_{i=1}^{q} \beta_i v^i.$$

# 1.3.2 凸函数

1、定义：设$F \subseteq R^n$为非空凸集，给定函数$f: F \to R$,

(1) 若对任意的$x, y \in F$及任意的$\alpha \in [0,1]$，有

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$$

则称函数$f$为凸集$F$上的凸函数。

若对任意的$x, y \in F$及任意的$\alpha \in [0,1]$，有

$$f(\alpha x + (1-\alpha)y) \ge \alpha f(x) + (1-\alpha) f(y)$$

则称函数$f$为凸集$F$上的凹函数.

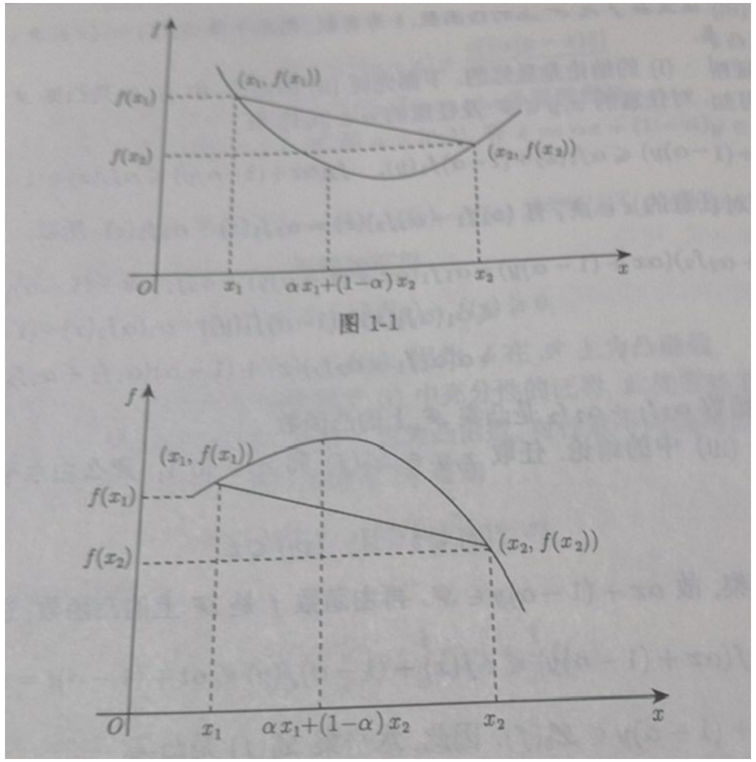


图 1-1

例. 给定向量$c \in R^n$.证明线性函数

$f(x) = c^T x$既是凸函数又是凹函数.

（2）若对任意的$x, y \in F$且$x \neq y$, 及任意的$\alpha \in (0,1)$有:

$$f(\alpha x+(1-\alpha)y) < \alpha f(x)+(1-\alpha)f(y)$$

则称函数$f$为凸集$F$上的严格凸函数。

（3）若存在常数$c>0$, 使得对任意的$x, y \in F$及任意的$\alpha \in (0,1)$有:

$$f(\alpha x+(1-\alpha)y) \leq \alpha f(x)+(1-\alpha)f(y)-c\alpha(1-\alpha)\|x-y\|^2$$

则称函数$f$为凸集$F$上的强凸函数（一致凸函数）.

## 凸函数的性质

### 命题1.3.2

设$F \subseteq R^n$是凸集，则：

(1) 函数$f$是$F$上的（严格）凸函数$\Leftrightarrow -f$是$F$上的（严格）凹函数；

(2) 设函数$f_1, f_2$是$F$上的凸函数, 实数$\alpha_1, \alpha_2 \geq 0$，则函数$\alpha_1 f_1 + \alpha_2 f_2$是$F$上的凸函数；

(3) 设函数$f$是$F$上的凸函数, $t$为实数，则水平集

$$L_t(f) = \{x \in F : f(x) \leq t\}$$

是凸集.

**证明:**

(2) 由$f_1, f_2$是凸集$F$上的凸函数知，对$\forall x, y \in F, \forall \alpha \in [0,1]$，有

$$f_1(\alpha x+(1-\alpha)y) \le \alpha f_1(x)+(1-\alpha)f_1(y);$$

$$f_2(\alpha x+(1-\alpha)y) \le \alpha f_2(x)+(1-\alpha)f_2(y);$$

由$(\alpha_1 f_1+\alpha_2 f_2)(z)=\alpha_1 f_1(z)+\alpha_2 f_2(z)$得

$$(\alpha_1 f_1+\alpha_2 f_2)(\alpha x+(1-\alpha)y)$$

$$=\alpha_1 f_1(\alpha x+(1-\alpha)y)+\alpha_2 f_2(\alpha x+(1-\alpha)y)$$

$$\le \alpha_1\alpha f_1(x)+\alpha_1(1-\alpha)f_1(y)+\alpha_2\alpha f_2(x)+\alpha_2(1-\alpha)f(y)$$

$$=\alpha(\alpha_1 f_1+\alpha_2 f_2)(x)+(1-\alpha)(\alpha_1 f_1+\alpha_2 f_2)(y)$$

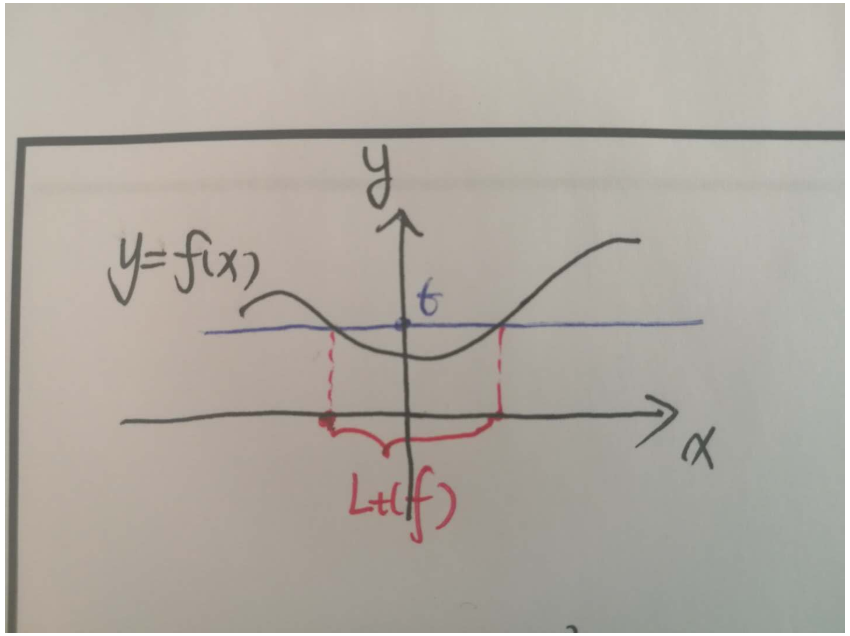
**证明:**

(3)$\forall x,y\in L_t(f),\forall\alpha\in[0,1],f(x)\le t,f(y)\le t.$

由$F$为凸集，$\alpha x+(1-\alpha)y\in F$

$f(\alpha x+(1-\alpha)y)\le\alpha f(x)+(1-\alpha)f(y)\le t$

所以$L_t(f)$为凸集。



## 可微凸函数的两个判别准则

定理1.3.4 （一阶判别定理）

设函数$f$在凸集$F \subseteq R^n$上可微，则

(1) $f$在$F$上为凸函数 $\Leftrightarrow$ 对任意的$x, y \in F$，有

$$f(y) - f(x) \geq \nabla f(x)^T (y - x).$$

(2) $f$在$F$上为严格凸函数 $\Leftrightarrow$ 对任意不同的$x, y \in F$, 有

$$f(y) - f(x) > \nabla f(x)^T (y - x).$$

**证明：　（定理 1.3.4）**

（1）$\Rightarrow \forall x, y \in F, \forall \alpha \in [0,1]$

一方面，$f(\alpha y + (1-\alpha)x) \le \alpha f(y) + (1-\alpha) f(x)$,即

$f(x + \alpha(y-x)) \le f(x) + \alpha[f(y) - f(x)]$.

另一方面，由一阶泰勒展开式，得

$f(x + \alpha(y-x)) = f(x) + \alpha \nabla f(x)^T (y-x) + o(\|\alpha(y-x)\|)$.

所以 $f(y) - f(x) \ge \nabla f(x)^T (y-x) + \dfrac{o(\|\alpha(y-x)\|)}{\alpha}$.

令$\alpha \to 0$，得 $f(y) - f(x) \ge \nabla f(x)^T (y-x)$.

充分性：

$\Leftarrow \forall x, y \in F, \forall \alpha \in [0,1]$，有 $z := \alpha x + (1-\alpha) y \in F$.

$$f(x) - f(z) \geq \nabla f(z)^T (x - z);$$

$$f(y) - f(z) \geq \nabla f(z)^T (y - z);$$

两式分别乘$\alpha$，（$1-\alpha$）并相加得：$\alpha f(x) + (1-\alpha) f(y) - f(z) \geq 0$,

即 $f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$.

因此 $f$在$F$上是凸函数.

**（2）充分性的证明类似于（1）的证明，下面只证明必要性。**

"$\Rightarrow$" $f(x)$为严格凸函数则$f(x)$必为凸函数，从而

$\forall x, y \in F$且$x \neq y$, 令$z = \frac{1}{2}(x+\mathrm{y})$, 则$z \in F$, 且

$f(z) - f(x) \geq \nabla f(x)^T (z - x)$;（由（1）的结论）

$f(z) = f(\frac{1}{2}x + \frac{1}{2}\mathrm{y}) < \frac{1}{2}f(x) + \frac{1}{2}f(\mathrm{y})$.（由$f$严格凸）

所以$\frac{1}{2}f(x) + \frac{1}{2}f(\mathrm{y}) > f(x) + \nabla f(x)^T (z - x)$.

整理得 $f(y) - f(x) > \nabla f(x)^T (y - x)$.

定理1.3.5 （二阶判别定理）

设函数 $f$ 在开凸集 $F \subseteq R^n$ 内二阶可微，则

(1) $f$ 在 $F$ 内为凸函数 $\Leftrightarrow$ 任意的 $x \in F$，$\nabla^2 f(x)$ 半正定.

(2) 若 $\forall x \in F$，$\nabla^2 f(x)$ 正定，则 $f$ 在 $F$ 内为严格凸函数.

**证明：定理1.3.5**

$(1) \Rightarrow \forall x \in F, 0 \neq y \in R^n$,由$F$是开集可知存在$\varepsilon > 0$,使得

$\forall \alpha \in (-\varepsilon, \varepsilon), x + \alpha y \in F$.

由定理1.3.4, $f(x+\alpha y) \geq f(x) + \alpha \nabla f(x)^T y$.

由$Taylor$展式得，

$$f(x+\alpha y) = f(x) + \alpha \nabla f(x)^T y + \frac{1}{2}\alpha^2 y^T \nabla^2 f(x) y + o(\|\alpha y\|^2).$$

所以$y^T \nabla^2 f(x) y + \dfrac{o(\|\alpha y\|^2)}{2\alpha^2} \geq 0$.

令$\alpha \to 0$，得 $y^T \nabla^2 f(x) y \geq 0$.由$y$的任意性得$\nabla^2 f(x)$正半定.

**证明：定理1.3.5**

(1) $\Leftarrow \forall x, y \in F$, 由Taloy展开式

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(\xi)(y-x)$$

其中$\xi = x + t(y-x), t \in (0,1)$.

$$\geq f(x) + \nabla f(x)^T (y-x),$$（由$\nabla^2 f(x)$对称半正定）

因此 $f$ 在$F$上是凸函数.（由定理1.3.4一阶判定条件）

定理1.3.6 （强凸函数的判定定理）

设 $f: R^n \to R$ 二次连续可微，则

$f$是强凸函数 $\Leftrightarrow \nabla^2 f(x)$一致正定.

注：$\nabla^2 f(x)$一致正定：

即存在常数$c>0$, 使得 $d^T \nabla^2 f(x) d \geq c\|d\|^2, (\forall x, d \in R^n)$

备注：设$\boldsymbol{A}$为$n$阶实对称矩阵，则

$\lambda_{\max} d^T d \geq d^T \boldsymbol{A} d \geq \lambda_{\min} d^T d,\ (\forall d \neq 0, d \in R^n)$

其中$\lambda_{\max}, \lambda_{\min}$分别是$\boldsymbol{A}$的最大、最小特征值.

# 1.3.3 凸规划

1. 定义：设$F \subseteq R^n$为凸集，$f: F \to R$为凸函数，则称

$$\min_{x \in F} f(x) \text{ 为凸规划问题.}$$

２ 凸规划的性质：

定理 1.3.7

(1) 凸规划问题的任一局部最优解 $x^*$ 为其全局最优解；

(2) 凸规划问题的最优解集 $S$ 为凸集；

(3) 若函数 $f$ 为非空凸集 $F$ 上的严格凸函数，且凸规划问题存在全局最优解，则其全局最优解唯一.

**证明:**

(1)(反证法) 假设$x^*$是凸规划问题的局部但非全局最优解，则至少存在一个$y^* \in F$使得$f(y^*) < f(x^*)$.

因为函数$f$为凸函数，$F$为凸集，则对$\forall \alpha \in [0,1]$,有

$$\begin{cases} \alpha y^* + (1-\alpha)x^* \in F; \\ f(\alpha y^* + (1-\alpha)x^*) \le \alpha f(y^*) + (1-\alpha) f(x^*) < f(x^*). \end{cases}$$

当$\alpha \to 0^+$时，$\alpha y^* + (1-\alpha)x^* \to x^*$.

因此在充分靠近$x^*$时，$f(\alpha y^* + (1-\alpha)x^*) < f(x^*)$,

这与$x^*$为局部最优解矛盾，故假设不成立.

**证明:**

(2)由于空集与单元素集合为凸集，不妨设凸规划问题的最优解集$S$至少含两个元素.

假设$x^*, y^* \in S$,则 $x^*, y^* \in F$，且$f(y^*) = f(x^*)$.

对$\forall \alpha \in [0,1]$，有

$f(x^*) \le f(\alpha x^* + (1-\alpha)y^*) \le \alpha f(x^*) + (1-\alpha) f(y^*) = f(x^*) = \min_{x\in F} f(x)$.

因此$f(\alpha x^* + (1-\alpha)y^*) = f(x^*) = \min_{x\in F} f(x)$,

从而 $\alpha x^* + (1-\alpha)y^* \in S$，故$S$是凸集.

(3)(反证法)假设全局最优点不唯一，则$\exists x^*, y^* \in S$且 $x^* \neq y^*$.

显然$x^*, y^* \in F$且$f(x^*) = f(y^*)$.

$\forall \alpha \in [0,1], \alpha x^* + (1-\alpha) y^* \in S$且

$f(\alpha x^* + (1-\alpha) y^*) < \alpha f(x^*) + (1-\alpha) f(y^*) = f(x^*)$

这与$x^*$为全局最优点矛盾，所以凸规划问题有唯一最优解。